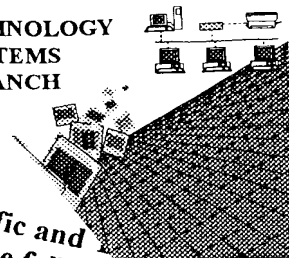


RAW SEQUENCE LISTING ERROR REPORT

TECHNOLOGY
SYSTEMS
BRANCH



The Biotechnology Systems Branch of the Scientific and
Center (STIC) detected errors when processing the following
form:

Application Serial Number: 09/719,533

Source:

Date Processed by STIC: PCT/09
9/10/2001

Information
readable

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.
PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:
1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE
APPLICANT, WITH A NOTICE TO COMPLY or,
2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A
NOTICE TO COMPLY

FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.
FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.
PATENTIN 2.1 e-mail help: patin21help@uspto.gov or phone 703-306-4119 (R. Wax)
PATENTIN 3.0 e-mail help: patin3help@uspto.gov or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER
VERSION 3.0 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND
TRADEMARK OFFICE WEBSITE. SEE BELOW:

Checker Version 3.0

The Checker Version 3.0 application is a state-of-the-art Windows based software program
employing a logical and intuitive user-interface to check whether a sequence listing is in
compliance with format and content rules. Checker Version 3.0 works for sequence listings
generated for the original version of 37 CFR §§1.821 - 1.825 effective October 1, 1990 (old
rules) and the revised version (new rules) effective July 1, 1998 as well as World Intellectual
Property Organization (WIPO) Standard ST.25.
Checker Version 3.0 replaces the previous DOS-based version of Checker, and is Y2K-
compliant. Checker allows public users to check sequence listings in Computer Readable form
(CRF) before submitting them to the United States Patent and Trademark Office (USPTO).
Use of Checker prior to filing the sequence listing is expected to result in fewer errored sequence
listings, thus saving time and money.

Checker Version 3.0 can be down loaded from the USPTO website at the following address:
<http://www.uspto.gov/web/offices/pac/checker>

Raw Sequence Listing Error Summary

ERROR DETECTED

SUGGESTED CORRECTION

SERIAL NUMBER: 09/27/533

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

1. Wrapped Nucleics
Wrapped Aminos
The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."
2. Invalid Line Length
The rules require that a line not exceed 72 characters in length. This includes white spaces.
3. Misaligned Amino
Numbering
The numbering under each 5th amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.
4. Non-ASCII
The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.
5. Variable Length
Sequence(s) _____ contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>.<223> section that some may be missing.
6. PatentIn 2.0
"bug"
A "bug" in PatentIn version 2.0 has caused the <220>.<223> section to be missing from amino acid sequences(s) _____. Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>.<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>.<223> sections for Artificial or Unknown sequences.
7. Skipped Sequences
(OLD RULES)
Sequence(s) _____ missing. If intentional, please insert the following lines for each skipped sequence:
(2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
(i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)
(xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
This sequence is intentionally skipped
Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.
8. Skipped Sequences
(NEW RULES)
Sequence(s) _____ missing. If intentional, please insert the following lines for each skipped sequence.
<210> sequence id number
<400> sequence id number
000
9. Use of n's or Xaa's
(NEW RULES)
Use of n's and/or Xaa's have been detected in the Sequence Listing.
Per 1.823 of Sequence Rules, use of <220>.<223> is MANDATORY if n's or Xaa's are present.
In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.
10. Invalid <213>
Response
Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>.<223> section is required when <213> response is Unknown or is Artificial Sequence
11. Use of <220>
Sequence(s) _____ missing the <220> "Feature" and associated numeric identifiers and responses.
Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.
(See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)
12. PatentIn 2.0
"bug"
Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.
13. Misuse of n
n can only be used to represent a single nucleotide in a nucleic acid sequence. N is not used to represent any value not specifically a nucleotide.

OIPE

RAW SEQUENCE LISTING

DATE: 09/10/2001

PATENT APPLICATION: US/09/719,533

TIME: 11:52:58

Input Set : A:\719533sequence.txt

Output Set: N:\CRF3\09102001\I719533.raw

Does Not Copy
Correct Diskette Needed

SEQUENCE LISTING

3 (1) GENERAL INFORMATION:
 5 (i) APPLICANT: Gon, Cheng Jin
 6 Lim, Gek Keow
 7 Leong, Ai Lin
 8 Zhao, Yi
 9 Chen, Wei Ning
 11 (ii) TITLE OF INVENTION: A VACCINE-INDUCED HEPATITIS B VIRAL
 12 STRAIN AND USES THEREOF
 14 (iii) NUMBER OF SEQUENCES: 11
 16 (iv) CORRESPONDENCE ADDRESS:
 17 (A) ADDRESSEE: Ladas & Parry
 18 (B) STREET: 26 West 61 Street
 19 (C) CITY: New York
 20 (D) STATE: New York
 21 (E) COUNTRY: USA
 22 (F) ZIP: 10023
 24 (v) COMPUTER READABLE FORM:
 25 (A) MEDIUM TYPE: Floppy disk
 26 (B) COMPUTER: IBM PC compatible
 27 (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 28 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
 30 (vi) CURRENT APPLICATION DATA:
 31 (A) APPLICATION NUMBER: US/09/719,533
 32 (B) FILING DATE: 10-Jul-2001
 33 (C) CLASSIFICATION: 435
 35 (vii) PRIOR APPLICATION DATA:
 36 (A) APPLICATION NUMBER: PCT/SG98/00045
 37 (B) FILING DATE: 19-JUN-1998
 39 (viii) ATTORNEY/AGENT INFORMATION:
 40 (A) NAME: Mass, Clifford J.
 41 (B) REGISTRATION NUMBER: 30,036
 42 (C) REFERENCE/DOCKET NUMBER: U-013108-9
 44 (ix) TELECOMMUNICATION INFORMATION:
 45 (A) TELEPHONE: (212) 708-1890

ERRORED SEQUENCES

47 (2) INFORMATION FOR SEQ ID NO: 1:
 48 (i) SEQUENCE CHARACTERISTICS:
 50 (A) LENGTH: 3215 base pairs
 51 (B) TYPE: nucleic acid
 52 (C) STRANDEDNESS: double
 53 (D) TOPOLOGY: circular
 55 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
 E--> 57 CTCCACCACT TTCCACCAAA CTCTTCAAGA TCCCAGAGTC AGGGCCCTGT

*See record
on Enver
summary sheet*

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/719,533

DATE: 09/10/2001

TIME: 11:52:58

Input Set : A:\719533sequence.txt

Output Set: N:\CRF3\09102001\I719533.raw

```

58 ACTTTCTCTGC      60
E--> 60 TGGTGGCTCC AGTTCAGGAA CAGTGAGCCC TGCTCAGAAT ACTGTCTCTG
61 CCATATCGTC      120
E--> 63 AATCTTATCG AAGACTGGGG ACCCTGTACC GAACATGGAG AACATCGCAT
64 CAGGACTCCT      180
E--> 66 AGGACCCCTG CTCGTGTTAC AGGCGGGGTT TTTCTTGTTG ACAAAAATCC
67 TCACAATACC      240
E--> 69 GCAGAGTCTA GACTCGTGGT GGACTTCTCT CAATTTTCTA GGGGGAACAC
70 CCGTGTGTCT      300
E--> 72 TGGCCAAAAT TCGCAGTCCC AAATCTCCAG TCACTCACCA ACCTGTTGTC
73 CTCGAATTTG      360
E--> 75 TCCTGGTTAT CGCTGGATGT GTCTGCGGCG TTTTATCATC TTCCTCTGCA
76 TCCTGCTGCT      420
E--> 78 ATGCCTCATC TTCTTGTTGG TTCTTCTGGA CTATCAAGGT ATGTTGCCCC
79 TTTGTCTCT      480
E--> 81 AATTCCAGGA TCAACAACAA CCAGCACCGG ACCATGCAAA ACCTGCACAA
82 CTCCTGCTCA      540
E--> 84 AGGAACCTCT ATGTTTCCCT CATGTTGCTG TACAAAACCT ACGGACAGAA
85 ACTGCACCTG      600
E--> 87 TATTCCTATC CCATCATCTT GGGCTTTTCGC AAAATACCTA TGGGAGTGGG
88 CTCAGTCCG      660
E--> 90 TTCTCTTTGG CTCAGTTTAC TAGTGCCATT TGTTCACTGG TTCGTAGGGC
91 TTTCCTCTCT      720
E--> 93 TGTCTGGCTT TCAGTTATAT GGATGATGTG GTTTTGGGGG CCAAGTCTGT
94 ACAACATCTT      780
E--> 96 GAGTCCCTTT ATGCCGCTGT TACCAATTTT CTTTGTCTT TGGGTATACA
97 TTTAAACCTT      840
E--> 99 CACAAAACAA AAAGATGGGG ATATTCCTCT AACTTCATGG GATATGTCAT
100 TGGGAGTTGG      900
E--> 102 GGCACATTGC CACAGGAACA TATTGTACAA AAAATCAAAA TGTGTTTTAG
103 GAAACTTCTT      960
E--> 105 GTAAACAGGC CTATTGATTG GAAAGTATGT CAACGAATTG TGGGTCTTTT
106 GGGGTTTGCC      1020
E--> 108 GCCCCTTTCA CGCAATGTGG ATATCCTGCT TTAATGCCTT TATATGCATG
109 TATACAAGCA      1080
E--> 111 AAACAGGCTT TACTTTCTC GCAAACCTAC AAGACCTTTC TAAGTAAACA
112 GATCTGAAAC      1140
E--> 114 CTTTACCCCG TTGCTCGGCA ACGCCCTGGT CTGTGCCAAG TGTGCTGA
115 CGCAACCCCG      1200
E--> 117 ACTGGTTGGG GCTTGGCCAT AGGCCATCAG CGCATGCGTG GAACCTTTGT
118 GTCTCTCTCT      1260
E--> 120 CCGATCCATA CTGCGGAACT CCTAGCCGCT TGTTTTGCTC GCAGCAGGTC
121 TGGGGCAAAA      1320
E--> 123 CTCATCGGGA CTGACAATTC TGTCGTGCTC TCCCGCAAGT ATACATCATT
124 TCCATGGCTG      1380
E--> 126 CTAGGCTGTG CTGCCAATG GATCCTGCGC GGGACGTCCT TTGTTTACGT
127 CCGCTCGGCG      1440
E--> 129 CTGAATCCCG CGGACGACCC CTCCCGGGGC CGCTTGGGGC TCTACCGCCC
130 GCTTCTCCCG      1500

```

*format
error
section 1*

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/719,533

DATE: 09/10/2001

TIME: 11:52:58

Input Set : A:\719533sequence.txt

Output Set: N:\CRF3\09102001\I719533.raw

205 CAAGGACAAC 3000
 E--> 207 TGGCCGGACG CCAACAAGGT GGGAGTGGGA GCATTCGGGC CAGGGTTCAC
 208 CCCTCCTCAT 3060
 E--> 210 GGGGGACTGT TGGGGTGGAG CCCTCAGGCT CAGGGCCTAC TCACAACTGT
 211 GCCAGCAGCT 3120
 E--> 213 CCTCCTCCTG CCTCCACCAA TCGGCAGTCA GGAAGGCAGC CTACTCCCTT
 214 ATCTCCACCT 3180
 215 CTAAGGGACA CTCATCCTCA GGCCATGCAG TGGAA 3215
 579 (2) INFORMATION FOR SEQ ID NO: 6:
 581 (i) SEQUENCE CHARACTERISTICS:
 582 (A) LENGTH: (35) base pairs → 36 shown
 583 (B) TYPE: nucleic acid
 584 (C) STRANDEDNESS: single
 585 (D) TOPOLOGY: linear
 587 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:
 E--> 589 ATAAGCTTATG ECCCTATCTT ATCAAACTT CCGGA
 644 (2) INFORMATION FOR SEQ ID NO: 11:
 646 (i) SEQUENCE CHARACTERISTICS:
 647 (A) LENGTH: 27 base pairs
 648 (B) TYPE: nucleic acid
 649 (C) STRANDEDNESS: single
 650 (D) TOPOLOGY: linear
 652 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:
 E--> 654 CTAAGCTTAG TTTCCGGAAG TGTTGAT

(35) 36

27 ← insert cumulative base
 total at right
 margin of each line

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/719,533

DATE: 09/10/2001

TIME: 11:53:00

Input Set : A:\719533sequence.txt

Output Set: N:\CRF3\09102001\I719533.raw

L:31 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:32 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
L:57 M:254 E: No. of Bases conflict, Input:0 Counted:50 SEQ:1
M:254 Repeated in SeqNo=1
L:589 M:254 E: No. of Bases conflict, Input:35 Counted:36 SEQ:6
L:589 M:204 E: No. of Bases differ, LENGTH:Input:35 Counted:36 SEQ:6
L:654 M:254 E: No. of Bases conflict, Input:0 Counted:27 SEQ:11



Creation date: 10-16-2003

Indexing Officer: ATEKLY - ALEM TEKLAY

Team: OIPEScanning

Dossier: 09719533

Legal Date: 10-26-2001

No.	Doccode	Number of pages
1	CTMS	2

Total number of pages: 2

Remarks:

Order of re-scan issued on